

Docket No. AUS920030540US1

**METHOD AND APPARATUS FOR QUALIFYING COLLECTION OF
PERFORMANCE MONITORING EVENTS BY TYPES OF INTERRUPT WHEN
INTERRUPT OCCURS**

CROSS REFERENCE TO RELATED APPLICATIONS

The present invention is related to the following applications entitled "Method and Apparatus for Counting Instruction Execution and Data Accesses", serial no. _____, attorney docket no. AUS920030477US1, filed on September 30, 2003; "Method and Apparatus for Selectively Counting Instructions and Data Accesses", serial no. _____, attorney docket no. AUS920030478US1, filed on September 30, 2003; "Method and Apparatus for Generating Interrupts Upon Execution of Marked Instructions and Upon Access to Marked Memory Locations", serial no. _____, attorney docket no. AUS920030479US1, filed on September 30, 2003; "Method and Apparatus for Counting Data Accesses and Instruction Executions that Exceed a Threshold", serial no. _____, attorney docket no. AUS920030480US1, filed on September 30, 2003; "Method and Apparatus for Counting Execution of Specific Instructions and Accesses to Specific Data Locations", serial no. _____, attorney docket no. AUS920030481US1, filed on September 30, 2003; "Method and Apparatus for Debug Support for Individual Instructions and Memory Locations", serial no. _____, attorney docket no. AUS920030482US1, filed on September 30, 2003; "Method and Apparatus to Autonomically Select Instructions for Selective Counting", serial no. _____, attorney

Docket No. AUS920030540US1

docket no. AUS920030483US1, filed on September 30, 2003; "Method and Apparatus to Autonomically Count Instruction Execution for Applications", serial no. _____, attorney docket no. AUS920030484US1, filed on September 30, 2003; "Method and Apparatus to Autonomically Take an Exception on Specified Instructions", serial no. _____, attorney docket no. AUS920030485US1, filed on September 30, 2003; "Method and Apparatus to Autonomically Profile Applications", serial no. _____, attorney docket no. AUS920030486US1, filed on September 30, 2003; "Method and Apparatus for Counting Instruction and Memory Location Ranges", serial no. _____, attorney docket no. AUS920030487US1, filed on September 30, 2003; "Method and Apparatus for Counting Interrupts by Type", serial no. _____, attorney docket no. AUS920030541US1, filed on _____, and "Method and Apparatus for Providing Pre and Post Handlers for Recording Events", serial no. _____, attorney docket no. AUS920030543US1, filed on _____. All of the above related applications are assigned to the same assignee, and incorporated herein by reference.

BACKGROUND OF THE INVENTION

1. Technical Field:

The present invention relates generally to an improved data processing system and, in particular, to a method and system for monitoring performance of the processor in a data processing system when an interrupt occurs. Still more particularly, the present invention relates to a method, apparatus, and computer instructions for qualifying collection of performance monitoring events by types of interrupt when interrupt occurs.

2. Description of Related Art:

A typical data processing system utilizes processors to execute a set of instructions in order to perform a certain task, such as reading a specific character from the main memory. However, as the number of tasks required to be executed by the processor increases, the efficiency of the processor's access patterns to memory and the characteristics of such access become important factors for engineers who want to optimize the system.

Currently, the prior art contains mechanisms that can count occurrences of software-selectable events, such as cache misses, instructions executed, I/O data transfer requests, and the time a given process may take to execute within a data processing system. One such mechanism is a performance monitor. A performance monitor performs monitoring on selected characteristics to assist analysis of a system by determining a machine's state at a particular time. This analysis provides

Docket No. AUS920030540US1

information of how the processor is used when instructions are executed and its interaction with the main memory when data are stored. In addition, the performance monitor may provide the amount of time that has passed between events in a processing system. The performance monitor provides counts of events that may be used by engineers to analyze system performance. This analysis may cause application code changes such as possible relocation of branch instructions and memory accesses to further optimize the performance of a system. Moreover, data may be gathered by the performance monitor on how the processor accesses the data processing system's level 1 and level 2 cache, and main memory in order to identify performance bottlenecks that are specific to a hardware or software environment.

The performance monitor described above does not provide the ability to further qualify the events that occur for a particular type of interrupt. An interrupt occurs, for example, when a device, such as a mouse or keyboard, raises an interrupt signal to notify the processor that an event has occurred. When the processor accepts an interrupt request, the processor completes its current instruction and passes the control to an interrupt handler. The interrupt handler executes an interrupt service routine that is associated with the interrupt. An interrupt may also be caused by a specific machine language operation code, for example Motorola 68000's TRAP, a product from Motorola. In this case, an unexpected software condition such as divide by zero causes the processor to store the current state, store

Docket No. AUS920030540US1

identifying information about the particular interrupt and pass control to an interrupt handler that handles this unexpected software condition.

It would be advantageous to have an improved method, apparatus, and computer instructions for qualifying counting of events that occur while servicing interrupt requests based on the type of interrupt that has occurred.

Docket No. AUS920030540US1

SUMMARY OF THE INVENTION

The present invention provides a method, apparatus and computer instructions for monitoring performance of a data processing system by qualifying events by type of interrupt when an interrupt occurs in the processor of the data processing system. The performance monitoring unit of the present invention is programmed to count events that are qualified based on the type of interrupt that has occurred. When the performance monitoring unit detects a particular type of the interrupt, the performance monitoring unit collects events that occur for that interrupt type. The resulting event information is later presented to the user for performance analysis.

BRIEF DESCRIPTION OF THE DRAWINGS

The novel features believed characteristic of the invention are set forth in the appended claims. The invention itself, however, as well as a preferred mode of use, further objectives and advantages thereof, will best be understood by reference to the following detailed description of an illustrative embodiment when read in conjunction with the accompanying drawings, wherein:

Figure 1 is an exemplary block diagram of a data processing system in which the present invention may be implemented;

Figure 2 is an exemplary block diagram of a processor system for processing information according to a preferred embodiment of the present invention;

Figure 3 is an exemplary diagram illustrating components for counting events by interrupt type when interrupt occurs in accordance with a preferred embodiment of the present invention;

Figure 4 is an exemplary diagram illustrating components of the performance monitoring unit and interrupt unit in accordance with a preferred embodiment of the present invention; and

Figure 5 is a flowchart outlining an exemplary process of qualify counting of events based on type of interrupt when interrupt occurs is depicted in accordance with a preferred embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

A preferred embodiment of the present invention monitors performance of data processing system while executing a program on the data processing system. In one embodiment, a performance monitoring unit is programmed to count events (such as cache misses, clock cycles, or other events that occur during the execution and processing of an interrupt) and to qualify those counted events by the type of interrupt during which they occur, and (in some embodiments) the state of the interrupt during which they occur. The information gathered in the innovative system and method can be presented to software or to a user for analysis.

In a preferred embodiment, the present invention is implemented in a computer system. For example, the computer system can be a client or a server in a client-server environment of a network. With reference now to **Figure 1**, an exemplary block diagram of a data processing system is shown in which the present invention may be implemented. Client **100** is an example of a computer, in which code or instructions implementing the processes of the present invention may be located. Client **100** employs a peripheral component interconnect (PCI) local bus architecture. Although the depicted example employs a PCI bus, other bus architectures such as Accelerated Graphics Port (AGP) and Industry Standard Architecture (ISA) may be used. Processor **102** and main memory **104** are connected to PCI local bus **106** through PCI bridge **108**. PCI bridge **108** also may include an integrated memory controller and cache

Docket No. AUS920030540US1

memory for processor **102**. Additional connections to PCI local bus **106** may be made through direct component interconnection or through add-in boards. In the depicted example, local area network (LAN) adapter **110**, small computer system interface SCSI host bus adapter **112**, and expansion bus interface **114** are connected to PCI local bus **106** by direct component connection. In contrast, audio adapter **116**, graphics adapter **118**, and audio/video adapter **119** are connected to PCI local bus **106** by add-in boards inserted into expansion slots. Expansion bus interface **114** provides a connection for a keyboard and mouse adapter **120**, modem **122**, and additional memory **124**. SCSI host bus adapter **112** provides a connection for hard disk drive **126**, tape drive **128**, and CD-ROM drive **130**. Typical PCI local bus implementations will support three or four PCI expansion slots or add-in connectors.

An operating system runs on processor **102** and is used to coordinate and provide control of various components within data processing system **100** in **Figure 1**. The operating system may be a commercially available operating system such as Windows XP, which is available from Microsoft Corporation. An object oriented programming system such as Java may run in conjunction with the operating system and provides calls to the operating system from Java programs or applications executing on client **100**. "Java" is a trademark of Sun Microsystems, Inc. Instructions for the operating system, the object-oriented programming system, and applications or programs are located on storage devices, such as hard disk drive

Docket No. AUS920030540US1

126, and may be loaded into main memory **104** for execution by processor **102**.

Those of ordinary skill in the art will appreciate that the hardware in **Figure 1** may vary depending on the implementation. Other internal hardware or peripheral devices, such as flash read-only memory (ROM), equivalent nonvolatile memory, or optical disk drives and the like, may be used in addition to or in place of the hardware depicted in **Figure 1**. Also, the processes of the present invention may be applied to a multiprocessor data processing system.

For example, client **100**, if optionally configured as a network computer, may not include SCSI host bus adapter **112**, hard disk drive **126**, tape drive **128**, and CD-ROM **130**. In that case, the computer, to be properly called a client computer, includes some type of network communication interface, such as LAN adapter **110**, modem **122**, or the like. As another example, client **100** may be a stand-alone system configured to be bootable without relying on some type of network communication interface, whether or not client **100** comprises some type of network communication interface. As a further example, client **100** may be a personal digital assistant (PDA), which is configured with ROM and/or flash ROM to provide non-volatile memory for storing operating system files and/or user-generated data. The depicted example in **Figure 1** and above-described examples are not meant to imply architectural limitations.

The processes of the present invention are performed by processor **102** using computer implemented instructions,

Docket No. AUS920030540US1

which may be located in a memory such as, for example, main memory **104**, memory **124**, or in one or more peripheral devices **126-130**.

Turning next to **Figure 2**, an exemplary block diagram of a processor system for processing information is depicted in accordance with a preferred embodiment of the present invention. Processor **210** may be implemented as processor **102** in **Figure 1**.

In a preferred embodiment, processor **210** is a single integrated circuit superscalar microprocessor. Accordingly, as discussed further herein below, processor **210** includes various units, registers, buffers, memories, and other sections, all of which are formed by integrated circuitry. Also, in the preferred embodiment, processor **210** operates according to reduced instruction set computer ("RISC") techniques. As shown in **Figure 2**, system bus **211** is connected to a bus interface unit ("BIU") **212** of processor **210**. BIU **212** controls the transfer of information between processor **210** and system bus **211**.

BIU **212** is connected to an instruction cache **214** and to data cache **216** of processor **210**. Instruction cache **214** outputs instructions to sequencer unit **218**. In response to such instructions from instruction cache **214**, sequencer unit **218** selectively outputs instructions to other execution circuitry of processor **210**.

In addition to sequencer unit **218**, in the preferred embodiment, the execution circuitry of processor **210** includes multiple execution units, namely a branch unit **220**, a fixed-point unit A ("FXUA") **222**, a fixed-point

Docket No. AUS920030540US1

unit B ("FXUB") **224**, a complex fixed-point unit ("CFXU") **226**, a load/store unit ("LSU") **228**, and a floating-point unit ("FPU") **230**. FXUA **222**, FXUB **224**, CFXU **226**, and LSU **228** input their source operand information from general-purpose architectural registers ("GPRs") **232** and fixed-point rename buffers **234**. Moreover, FXUA **222** and FXUB **224** input a "carry bit" from a carry bit ("CA") register **239**. FXUA **222**, FXUB **224**, CFXU **226**, and LSU **228** output results (destination operand information) of their operations for storage at selected entries in fixed-point rename buffers **234**. Also, CFXU **226** inputs and outputs source operand information and destination operand information to and from special-purpose register processing unit ("SPR unit") **237**.

FPU **230** inputs its source operand information from floating-point architectural registers ("FPRs") **236** and floating-point rename buffers **238**. FPU **230** outputs results (destination operand information) of its operation for storage at selected entries in floating-point rename buffers **238**.

In response to a Load instruction, LSU **228** inputs information from data cache **216** and copies such information to selected ones of rename buffers **234** and **238**. If such information is not stored in data cache **216**, then data cache **216** inputs (through BIU **212** and system bus **211**) such information from a system memory **260** connected to system bus **211**. Moreover, data cache **216** is able to output (through BIU **212** and system bus **211**) information from data cache **216** to system memory **260** connected to system bus **211**. In response to a Store

Docket No. AUS920030540US1

instruction, LSU **228** inputs information from a selected one of GPRs **232** and FPRs **236** and copies such information to data cache **216**.

Sequencer unit **218** inputs and outputs information to and from GPRs **232** and FPRs **236**. From sequencer unit **218**, branch unit **220** inputs instructions and signals indicating a present state of processor **210**. In response to such instructions and signals, branch unit **220** outputs (to sequencer unit **218**) signals indicating suitable memory addresses storing a sequence of instructions for execution by processor **210**. In response to such signals from branch unit **220**, sequencer unit **218** inputs the indicated sequence of instructions from instruction cache **214**. If one or more of the sequence of instructions is not stored in instruction cache **214**, then instruction cache **214** inputs (through BIU **212** and system bus **211**) such instructions from system memory **260** connected to system bus **211**.

In response to the instructions input from instruction cache **214**, sequencer unit **218** selectively dispatches the instructions to selected ones of execution units **220**, **222**, **224**, **226**, **228**, and **230**. Each execution unit executes one or more instructions of a particular class of instructions. For example, FXUA **222** and FXUB **224** execute a first class of fixed-point mathematical operations on source operands, such as addition, subtraction, ANDing, ORing and XORing. CFXU **226** executes a second class of fixed-point operations on source operands, such as fixed-point multiplication and division. FPU **230** executes floating-point operations on

Docket No. AUS920030540US1

source operands, such as floating-point multiplication and division.

As information is stored at a selected one of rename buffers **234**, such information is associated with a storage location (e.g. one of GPRs **232** or carry bit (CA) register **242**) as specified by the instruction for which the selected rename buffer is allocated. Information stored at a selected one of rename buffers **234** is copied to its associated one of GPRs **232** (or CA register **242**) in response to signals from sequencer unit **218**. Sequencer unit **218** directs such copying of information stored at a selected one of rename buffers **234** in response to "completing" the instruction that generated the information. Such copying is called "writeback."

As information is stored at a selected one of rename buffers **238**, such information is associated with one of FPRs **236**. Information stored at a selected one of rename buffers **238** is copied to its associated one of FPRs **236** in response to signals from sequencer unit **218**. Sequencer unit **218** directs such copying of information stored at a selected one of rename buffers **238** in response to "completing" the instruction that generated the information.

Processor **210** achieves high performance by processing multiple instructions simultaneously at various ones of execution units **220**, **222**, **224**, **226**, **228**, and **230**. Accordingly, each instruction is processed as a sequence of stages, each being executable in parallel with stages of other instructions. Such a technique is called "pipelining." In a significant aspect of the

Docket No. AUS920030540US1

illustrative embodiment, an instruction is normally processed as six stages, namely fetch, decode, dispatch, execute, completion, and writeback.

In the fetch stage, sequencer unit **218** selectively inputs (from instruction cache **214**) one or more instructions from one or more memory addresses storing the sequence of instructions discussed further hereinabove in connection with branch unit **220**, and sequencer unit **218**.

In the decode stage, sequencer unit **218** decodes up to four fetched instructions.

In the dispatch stage, sequencer unit **218** selectively dispatches up to four decoded instructions to selected (in response to the decoding in the decode stage) ones of execution units **220**, **222**, **224**, **226**, **228**, and **230** after reserving rename buffer entries for the dispatched instructions' results (destination operand information). In the dispatch stage, operand information is supplied to the selected execution units for dispatched instructions. Processor **210** dispatches instructions in order of their programmed sequence.

In the execute stage, execution units execute their dispatched instructions and output results (destination operand information) of their operations for storage at selected entries in rename buffers **234** and rename buffers **238** as discussed further hereinabove. In this manner, processor **210** is able to execute instructions out-of-order relative to their programmed sequence.

In the completion stage, sequencer unit **218** indicates an instruction is "complete." Processor **210**

Docket No. AUS920030540US1

"completes" instructions in order of their programmed sequence.

In the writeback stage, sequencer **218** directs the copying of information from rename buffers **234** and **238** to GPRs **232** and FPRs **236**, respectively. Sequencer unit **218** directs such copying of information stored at a selected rename buffer. Likewise, in the writeback stage of a particular instruction, processor **210** updates its architectural states in response to the particular instruction. Processor **210** processes the respective "writeback" stages of instructions in order of their programmed sequence. Processor **210** advantageously merges an instruction's completion stage and writeback stage in specified situations.

In the illustrative embodiment, each instruction requires one machine cycle to complete each of the stages of instruction processing. Nevertheless, some instructions (e.g., complex fixed-point instructions executed by CFXU **226**) may require more than one cycle. Accordingly, a variable delay may occur between a particular instruction's execution and completion stages in response to the variation in time required for completion of preceding instructions.

Completion buffer **248** is provided within sequencer **218** to track the completion of the multiple instructions which are being executed within the execution units. Upon an indication that an instruction or a group of instructions have been completed successfully, in an application specified sequential order, completion buffer **248** may be utilized to initiate the transfer of the

Docket No. AUS920030540US1

results of those completed instructions to the associated general-purpose registers.

In addition, processor **210** also includes performance monitor unit **240**, which is connected to instruction cache **214** as well as other units in processor **210**. Operation of processor **210** can be monitored utilizing performance monitor unit **240**, which in this illustrative embodiment is a software-accessible mechanism capable of providing detailed information descriptive of the utilization of instruction execution resources and storage control. Although not illustrated in **Figure 2**, performance monitor unit **240** is coupled to each functional unit of processor **210** to permit the monitoring of all aspects of the operation of processor **210**, including, for example, reconstructing the relationship between events, identifying false triggering, identifying performance bottlenecks, monitoring pipeline stalls, monitoring idle processor cycles, determining dispatch efficiency, determining branch efficiency, determining the performance penalty of misaligned data accesses, identifying the frequency of execution of serialization instructions, identifying inhibited interrupts, and determining performance efficiency. The events of interest also may include, for example, time for instruction decode, execution of instructions, branch events, cache misses, and cache hits.

Performance monitor unit **240** includes an implementation-dependent number (e.g., 2-8) of counters **241-242**, labeled PMC1 and PMC2, which are utilized to count occurrences of selected events. Performance monitor

Docket No. AUS920030540US1

unit **240** further includes at least one monitor mode control register (MMCR). In this example, two control registers, MMCRs **243** and **244** are present that specify the function of counters **241-242**. Counters **241-242** and MMCRs **243-244** are preferably implemented as SPRs that are accessible for read or write via MFSPR (move from SPR) and MTSPR (move to SPR) instructions executable by CFXU **226**. However, in one alternative embodiment, counters **241-242** and MMCRs **243-244** may be implemented simply as addresses in I/O space. In another alternative embodiment, the control registers and counters may be accessed indirectly via an index register. This embodiment is implemented in the IA-64 architecture in processors from Intel Corporation.

Additionally, processor **210** also includes interrupt unit **250**, which is connected to instruction cache **214**. Additionally, although not shown in **Figure 2**, interrupt unit **250** is connected to other functional units within processor **210**. Interrupt unit **250** may receive signals from other functional units and initiate an action, such as starting an error handling or trap process. In these examples, interrupt unit **250** is employed to generate interrupts and exceptions that may occur during execution of a program. In addition, for each performance monitoring counter such as PMC1 **241**, there is a corresponding interrupt unit control register (IUCR), such as IUCR2 **252** that is accessible by the interrupt unit **250**. In this example, IUCR1 **252** corresponds to PMC1 **241**. IUCR2 **252** corresponds to PMC2 **242**. IUCR1 **252** and IUCR2 **254** identify which interrupt generates a signal, so

Docket No. AUS920030540US1

that PMC1 **241** and PMC2 **242** may count events associated with that interrupt.

The present invention provides a method, apparatus, and computer instructions for monitoring performance of a data processing system such as client **100** in **Figure 1** by qualifying events by type of interrupt when an interrupt occurs in the processor of the data processing system. The processor of the data processing system may be implemented as processor **210** in **Figure 2**.

In a preferred embodiment, the present invention provides a performance monitoring unit (PMU), such as performance monitoring unit **240** in **Figure 2**, that is programmable using counter control registers (CCR) such as MMCR **243** and MMCR **244** to count events that occur during execution of an interrupt. The counted events may be qualified by type of interrupt during which they occur. These events are currently known to the performance monitoring units, for example, instructions retired, TLB misses, level 2 cache misses, or clock cycles etc. In a preferred embodiment, the innovative performance monitor unit allows users to qualify events or conditions by interrupt type, preferably using hardware counters and registers of the performance monitoring unit. When an interrupt is signaled, the interrupt signal is used to identify whether the counting mechanism of the present invention is programmed to count events during that type of interrupt.

The counter control registers (CCR) in the performance monitoring unit may be extended to include an interrupt unit (IU) trace field, a interrupt unit (IU)

Docket No. AUS920030540US1

state field, and a interrupt unit (IU) state mask field. The IU trace field allows user to control counting by enabling or disabling tracing. The IU state field represents the current state of interrupt processing, for example, interrupt on, interrupt taken, interrupt acknowledged and interrupt off. The IU state mask field indicates the state of interrupt processing during which tracing is to occur. These fields are used by the performance monitoring unit to count events that occur during the various states of the interrupt, and to parse the counting by state. The resulting data may be read by the user or a trace application, for example, from the performance monitoring unit for performance analysis at a later time in order to better optimize the system.

In addition, the performance monitoring unit may count multiple events at the same time by using multiple performance monitoring counters. The mechanism of the present invention allows the performance monitoring unit to program each counter to count a particular event. For example, counter 1 may be programmed to count cache misses for interrupt vector 1, counter 2 may be programmed to count cycles for interrupt vector 2 etc. Preferred embodiments of the present invention do not limit the qualifications for the counters. The counters may be programmed to count multiple events based on the hardware available to the user.

Furthermore, the performance monitoring unit supports counting of events for nested interrupts. Nested interrupts occur when an interrupt occurs while another interrupt is active. The performance monitoring

Docket No. AUS920030540US1

unit of the present invention enables the performance monitoring counters to limit counting only to the on or off states of interrupt processing. This mechanism may be useful to engineers for investigating the efficiency of interrupt processing itself.

Turning next to **Figure 3**, an exemplary diagram illustrating components for counting events by interrupt type when an interrupt occurs is depicted in accordance with a preferred embodiment of the present invention. As depicted in **Figure 3**, in a preferred embodiment, the central processing unit (CPU) **302** may be implemented as processor **210** in **Figure 2**. When an interrupt occurs, the central processing unit (CPU) **302** sends a signal to interrupt unit **304** to request interrupt unit **304** generate an interrupt signal for a particular interrupt type.

Interrupt unit **304** may include a number of interrupt unit control registers (IUCRs) that are accessible to interrupt unit **304** based upon the number of performance monitoring counters in the performance monitoring unit **306**. There is a one to one relationship between the IUCR and the performance monitoring counters. In this example, there are two IUCRs, IUCR1 **308** that corresponds to performance monitoring counter PMC1 **312** and IUCR2 **310** that corresponds to performance monitoring counter PMC2 **314**. IUCR such as IUCR1 **308** includes a type field that indicates which interrupt type is to generate a performance monitoring counter signal. The type field in the IUCR is later examined by interrupt unit **304** to see if it is an interrupt type of interest, i.e., whether

Docket No. AUS920030540US1

events are to be counted during execution of the interrupt. If the interrupt type signal is an interrupt type of interest, interrupt unit **304** sends a Tracing On/Off signal to the performance monitoring unit **306** in order to enable or disable counting.

Also shown in **Figure 3**, Performance monitoring unit **306** includes an implementation-dependent number of hardware counters such as hardware counters **312** and **314**, also referred to herein as performance monitoring counters. PMC1 **312** and PMC2 **314** count events that occur during a particular state of interrupt processing. Performance monitoring unit **306** also includes a number of counter control registers (CCRs) such as CCR1 **316** and CCR2 **318**, which may be implemented as MCCR1 **233** and MCCR2 **234** in **Figure 2**. When CCR1 **316** and CCR2 **318** receive signals from interrupt unit **204** to enable or disable tracing events of an interrupt type and state changes during interrupt processing, CCR1 **316** and CCR2 **318** controls counting of the performance monitoring counters PMC1 **312** and PMC2 **314** by storing the current state of the interrupt and the state of the interrupt to trace, so that when tracing is enabled, performance monitoring counters PMC1 **312** and PMC2 **314** may count events of a particular interrupt state. The resulting trace data is then stored and may be collected at a later time from performance monitoring counters **312** and **314**.

With reference to **Figure 4**, an exemplary diagram illustrating components of the performance monitoring unit and interrupt unit is depicted in accordance with a preferred embodiment of the present invention. As

Docket No. AUS920030540US1

depicted in **Figure 4**, in a preferred embodiment, the performance monitoring unit may include a number of counter control registers (CCRs), which may be implemented as memory mode register (MMCR) such as MMCR1 **233** and MMCR2 **234** in **Figure 2**. The counter control registers control functions of the performance monitoring counters. The counter control register **402** may include 3 fields: an interrupt unit (IU) trace field **404**, an interrupt unit (IU) state field **406**, and an interrupt unit (IU) state mask field **408**. In this example, the IU trace field **402** is a one bit field, represented by bit 0; the IU state field **406** is a 2 bit field, represented by bit 1 and 2; and the IU state mask field **408** is a 3 bit field, represented by bit 3 to bit 5.

Also shown in **Figure 4**, trace **410** is an example of the IU trace field **404**. When the bit is 0, tracing by the performance monitoring counter is off or disabled. When bit is 1, tracing by the performance monitoring counter is on or enabled. Trace **410** is set by a Tracing On/Off signal that is raised by the interrupt unit once the interrupt unit examines the interrupt unit control register (IUCR) to determine whether interrupt type is of interest. For example, incoming interrupt signals can be compared to a stored set of interrupts that are designated for the counting process.

Next, an example IU state field **406** is represented by State **416**. State **416** is read only by the performance monitoring unit and is used to store the current state of the interrupt. When the interrupt state is interrupt on **418**, the bits in state **410** are set to 01. When the

Docket No. AUS920030540US1

interrupt state is interrupt taken **420**, the bits in state **410** are set to 10. When the interrupt state is interrupt acknowledged **422**, the bits in state **410** are set to 11. When the interrupt state is interrupt off **424**, the bits in state **410** are set to 00.

The last field in the counter control register is the IU state mask field **408**. State mask **426** represents an example of IU state mask field **408**. The IU state mask field **426** indicates during which state of the interrupt to count events. When the state during which to count is interrupt ON **428**, state mask bits **426** are set to 001. When the state during which to count is interrupt taken **430**, State mask **426** bits are set to 010. When the state during which to count is interrupt acknowledged, State mask bits **426** are set to 100.

Lastly, an interrupt unit control register (IUCR) **440** is shown in **Figure 4**. IUCR may be accessible by the interrupt unit. IUCR **440** includes a type field **442** that has a size of a byte (8 bits). The type field **442** represents which interrupt type is to generate a Tracing On/Off signal to the performance monitoring unit.

Preferred embodiments of the present invention can be used to count events based on the type of interrupt during which they occur. Interrupt types are noted in an IUCR field, which indicates the types of interrupt during which monitoring of events is to occur. When an interrupt occurs, the interrupt type is compared to entries in the IUCR, and if the occurring interrupt matches an entry in the IUCR, events that occur during that interrupt are to be counted. A signal is sent to the PMU to turn on

Docket No. AUS920030540US1

tracing, which causes the PMU to set the CCR to the state of the interrupt which is then occurring. As the interrupt goes through its states, the PMU continues to indicate what state the interrupt is in, so that events counted during the interrupt can be parsed by the interrupt state during which they occurred.

Turning to **Figure 5**, a flowchart outlining an exemplary process of qualify counting of events based on type of interrupt when interrupt occurs is depicted in accordance with a preferred embodiment of the present invention. As depicted in **Figure 5**, the process begins when supporting software (such as performance monitoring software) interfaces with the hardware to initiate interrupt monitoring (step **502**). The performance monitoring counter is then set by supporting software to monitor some event (step **504**), such as cache misses or clock cycles, for example. The supporting software also sets the interrupt unit control register (IUCR) type field to indicate which interrupt type is to be monitored (step **506**), i.e., during which interrupt types events will be counted. Next, the supporting software sets the interrupt unit (IU) trace field of the counter control register (CCR) that corresponds to a performance monitoring counter used to count events (step **508**). Step **508** enables tracing of events.

At some time after the IU trace field is set in step **508**, application code is executed for an application that is to be monitored for performance (step **510**). When an interrupt occurs during code execution (step **512**), the interrupt unit examines the IUCR type field (step **514**).

Docket No. AUS920030540US1

A determination is made as to whether the interrupt is the same type as the IUCR type, which is an interrupt type of interest (step **516**). If the interrupt type is not the same as the IUCR type, the process returns to step **512** to wait for the next interrupt to occur. If the interrupt type is the same as the IUCR type, the interrupt unit raises a Tracing On/Off signal and sends the signal to the performance monitoring unit (step **518**) in order to start tracing events for the interrupt. Once the signal is received, the performance monitoring unit sets the counter control register's state field to interrupt on by setting the bits to 01 (step **520**).

Next, the interrupt state changes to interrupt taken, the interrupt unit in turn raises a state change signal and sends the signal to the performance monitoring unit (step **522**). Consequently, the performance monitoring unit sets the counter control register's state field to interrupt taken by setting the bits to 10 (step **524**). Another interrupt state change causes the interrupt unit to raise a state change signal and sends the signal to the performance monitoring unit (step **526**).

Similarly, the performance monitoring unit sets the counter control register's state field to interrupt acknowledged by setting the bits to 11 (step **528**). When the interrupt handler finishes executing the interrupt service routine, it signals the interrupt unit an IRET (interrupt return) (step **530**). The interrupt unit raises a Tracing On/Off signal and sends the signal to the performance monitoring unit in order to stop tracing events for the interrupt (step **532**). (Note that

Docket No. AUS920030540US1

embodiments of the present invention can trace across one or a plurality of states of the interrupt, as described above.) Finally, the performance monitoring unit receives the signal and sets the counter control register state field to interrupt off by setting the bits to 00 (step 534). Thus, the process terminates thereafter.

Thus, the present invention provides a performance monitoring unit that allows the user to qualify events that occur by type of interrupt using one or more hardware counters. Hardware counters collect events that occur while servicing interrupt requests in order for a user to monitor performance of the system when a particular type of interrupt is being processed. Using a number of hardware counters of the present invention, a user may qualify different events for different interrupt type.

The performance monitoring unit also allows the user to qualify events that occur based upon the state of interrupt processing, so that a user may further collect events that occur during a particular interrupt state. Furthermore, the performance monitoring unit provides support for user to qualify events that occur when a nested interrupt occurs by using a counter that is associated with the nested interrupt. Moreover, the performance monitor unit may also count conditions such as clock cycles for a particular interrupt type at a particular interrupt state.

It is important to note that while the present invention has been described in the context of a fully functioning data processing system, those of ordinary

Docket No. AUS920030540US1

skill in the art will appreciate that the processes of the present invention are capable of being distributed in the form of a computer readable medium of instructions and a variety of forms and that the present invention applies equally regardless of the particular type of signal bearing media actually used to carry out the distribution. Examples of computer readable media include recordable-type media, such as a floppy disk, a hard disk drive, a RAM, CD-ROMs, DVD-ROMs, and transmission-type media, such as digital and analog communications links, wired or wireless communications links using transmission forms, such as, for example, radio frequency and light wave transmissions. The computer readable media may take the form of coded formats that are decoded for actual use in a particular data processing system.

The description of the present invention has been presented for purposes of illustration and description, and is not intended to be exhaustive or limited to the invention in the form disclosed. Many modifications and variations will be apparent to those of ordinary skill in the art. The embodiment was chosen and described in order to best explain the principles of the invention, the practical application, and to enable others of ordinary skill in the art to understand the invention for various embodiments with various modifications as are suited to the particular use contemplated.